DOCUMENT RESUME

ED 433 750 HE 032 300

AUTHOR	Abou-Sayf, Frank K.
TITLE	The Integrity of Ethnicity Data. AIR 1999 Annual Forum Paper.
PUB DATE	1999-06-00
NOTE	14p.; Paper presented at the Annual Forum of the Association for Institutional Research (39th, Seattle, WA, May 30-June 3, 1999).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	College Students; Community Colleges; *Data Collection; Ethnic Groups; Ethnic Origins; *Ethnicity; *Reliability; Student Surveys; Two Year College Students; Two Year Colleges
IDENTIFIERS	*AIR Forum; *University of Hawaii Kapiolani Community College

ABSTRACT

Since ethnicity is so often used in research on educational outcomes, this study examined the consistency with which students at Kapiolani Community College (Hawaii) reported their ethnic background. The study compared two different sources of ethnic data: the college application form and an interest inventory that all students fill out and submit with the application. A total of 3,993 records were examined. Analysis indicated that the percentage of inconsistency in reporting ethnicity ranged from a low of 16.2 percent in the case of Koreans to a high o 42.3 percent in the case of Hawaiians. Findings suggest that the more ethnically mixed students are the most likely to show discrepancies in reporting their ethnic background and that the specific way items on a questionnaire are phrased affect student choice. (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 433 750

THE INTEGRITY OF ETHNICITY DATA

Frank K. Abou-Sayf

Assistant to the Provost

and Director, Institutional Research

Kapiolani Community College

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
D. Vura
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

HE 032 300

AIR
for Management Research, Policy Analysis, and Planning

This paper was presented at the Thirty-Ninth Annual Forum of the Association for Institutional Research held in Seattle, Washington, May 30-June 3, 1999.
This paper was reviewed by the AIR Forum Publications Committee and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC Collection of AIR Forum Papers.

Dolores Vura
Editor
Air Forum Publications

THE INTEGRITY OF ETHNICITY DATA

Abstract

A study of self-reported student ethnicity at a two-year community college reveals a lack of consistency in these reports. The results indicate that the more ethnically mixed the students the more likely they are to be inconsistent in their reports. Another factor that influences consistency in reporting ethnicity is the format in which students are requested to report it. The implications resulting from these findings are discussed.

THE INTEGRITY OF ETHNICITY DATA

Introduction

In academe, one of the most frequent demographic breakdowns that are used to analyze cohorts' composition and educational outcomes is ethnicity. As people mingle and mix, aided by the ease of travel and migration, mixed ethnic backgrounds increase. The purpose of this investigation was to determine the consistency with which students report their ethnic background.

At Kapiolani Community College (KCC) in Honolulu, students are highly mixed ethnically, reflecting the society in which they belong. As such, KCC's database constitutes a good laboratory to investigate the hypothesis advanced above.

Literature Review

A review of the literature indicates that student report their demographic information in an inconsistent manner. Walsh (1967) compared student self-reported demographic information in interviews and questionnaires with that obtained from personnel records. He found discrepancies ranging from 2 percent to 52 percent, depending on the type of information. Laing, Sawyer, and Noble (1987) compared the consistency of self-reported extra-curricular activities with reports about the students available at their school, and found differences of about 10 percent. Porst and Zeifang (1987) investigated the

consistency with which students report demographic information by administering the same instrument twice within a one-month interval. They found discrepancy rates ranging from 0 to 64 percent. Takalkar, Waugh, and Micceri (1993) compared students' admission status to nine Florida universities with their self-reported application status. They found an overall error rate of 4.2 percent, with ranges from 3.6 percent to 24.6 percent.

Methodology

Two different sources of ethnic data were identified: the application form and an interest inventory that all students fill out and submit along with the application form. In these sources, three items in which students self-report their ethnicity were used to compare students' selections. These items are listed in Table 1 along with the source from which each was obtained. As shown, these items were obtained from two sources: one item from the Application Form, and two items, labeled (1) and (2) in the table, from the Interest Inventory.

Table 1

Sources of Ethnic Information

Source	*Item*
Application Form	*Which of these best describes your ethnic background (choose only one)?* - *Afro-American* - *Caucasian* - *Chinese* - *Filipino* - *Vietnamese* - *Japanese* - *Korean* - *Laotian* - *Samoan* - *None of the above*
Interest Inventory (1)	*Determine the **one** category which you believe **best represents** your ethnic background.* - *American Indian or Alaskan Native* - *Chinese* - *Filipino* - *Hawaiian or part Hawaiian* - *Korean* - *Japanese* - *Pacific Islander* - *Other Asian* - *Persons of the Indian Subcontinent: Afghanistan, Bangladesh, India, Nepal, Pakistan, Sri Lanka* - *Black* - *Puerto Rican* - *Other Hispanic* - *Mixed Hispanic* - *Caucasian, including North Africa and the Middle East, and excluding Spain and Portugal* - *Portuguese* - *Mixed*
Interest Inventory (2)	*Are you Hawaiian or Part Hawaiian (any ancestor living in Hawaii prior to the arrival of Captain Cook)?* - *Yes* - *No*

All students applying to KCC for fall 1994 and fall 1995 terms were selected. A total of 3,993 records were obtained. The ethnic background frequency distribution, as reported in the Interest Inventory item (1), is listed in Table 2.

Table 2

Ethnic Distribution

Ethnic Composition	*Frequency*	*Percent*
- American Indian or Alaskan Native	12	0.30
- Chinese	341	8.54
- Filipino	768	19.23
- Hawaiian or Part Hawaiian	393	9.84
- Korean	205	5.13
- Japanese	766	19.18
- Pacific Islander	101	2.53
- Other Asian	145	3.63
- Mixed Asian	112	2.80
- Persons of the Indian Subcontinent: Afghanistan, Bangladesh, India, Nepal, Pakistan, Sri Lanka	17	0.43
- African or African American	55	1.38
- Puerto Rican	25	0.63
- Other Hispanic	42	1.05
- Mixed Hispanic	18	0.45
- Caucasian, including North Africa and the Middle East, and excluding Spain and Portugal	646	16.18
- Portuguese	31	0.78
- Mixed	316	7.91
Total	***3,993***	***100.00***

After eliminating those ethnic groups with small sample sizes, seven categories remained. These are: Caucasian, Chinese, Filipino, Hawaiian or Part-Hawaiian, Japanese, Korean, and Mixed. The last category was then

eliminated due to the ambiguity that may result from using it for comparison purposes, leaving six categories for this study. For all of these groups except the Hawaiian and Part-Hawaiian category, the item from the Application Form and item (1) from the Interest Inventory were compared. For the Hawaiian and Part-Hawaiian category, the two items from the Interest Inventory were compared with each other. The reason for this change is that the Application Form item does not contain a choice for Hawaiian or Part-Hawaiian. For each group, the percent of consistent selections was then calculated as the number of students who selected the same ethnicity in both questions divided by the number of students who selected that ethnicity in at least one of the two items above.

Results

The results of this comparison are presented in Table 3 and illustrated in Figure 1.

Table 3

Inconsistent Responses

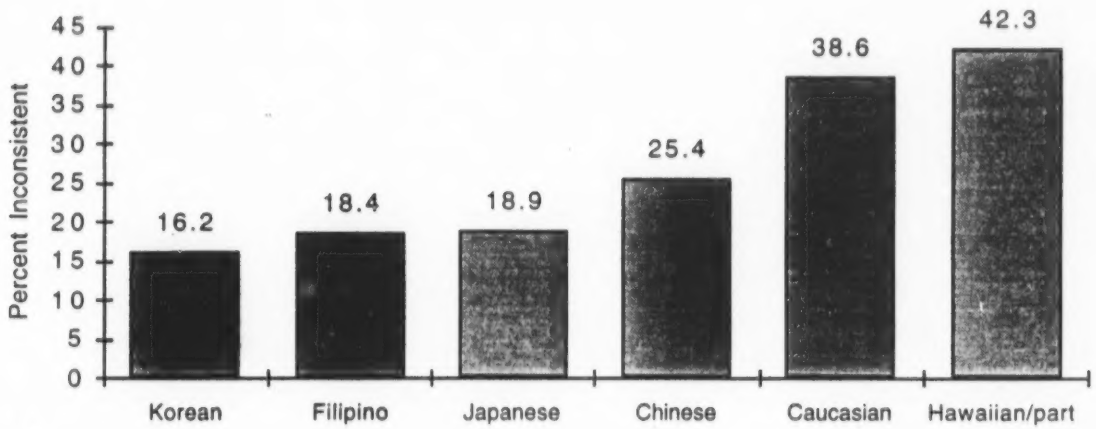
Ethnicity	*Percent Inconsistent*
Caucasian	38.6
Chinese	25.4
Filipino	18.4
Hawaiian/part	42.3
Japanese	18.9
Korean	16.2

Discussion

The data show that the percentage of inconsistency in reporting ethnicity ranged from a low of 16.2 in the case of Koreans to a high of 42.3 in the case of Caucasians. In this study, memory seems an unlikely source of discrepancy since the two instruments have to be submitted at the same time, and could be assumed to have been filled out shortly after one another. Although carelessness could not be ruled out as one reason for this discrepancy, it is unlikely to be the only factor. In this case, it was observed that the

Figure 1

Inconsistent Responses



percentages of inconsistent selections reflect to some extent the historical mixing of ethnic groups in Hawaii: Koreans, being the most recent immigrants to the islands among the groups being studied, are less likely to be racially mixed as, say, Caucasians or Hawaiians. It is no surprise that this ethnic group also reports the least amount of discrepancy. This finding suggests that the more ethnically mixed students are the most likely to show discrepancies in reporting their ethnic background.

Another factor that appears to play a determining role in the particular context of KCC is the way the items are phrased. For example, the Interest Inventory item (1) specifically includes people from North Africa and the Middle East as Caucasian, while specifically excluding people from Spain and Portugal. The Application Form item does not offer this breakdown and offers no alternative for those groups included and excluded from the Caucasian category in the Inventory, except perhaps the "None of the Above" category. Also, the Hispanic categories listed in the Interest Inventory item (1) are absent from the Application Form item. These two differences in statements are likely to have an impact on the number of students selecting the "Caucasian" category. Also, in case of the Hawaiian/Part-Hawaiian category, many students who are part Hawaiian could choose this category on item (1) of the Inventory, while answering No to the question of whether any of their ancestors were living in Hawaii prior to the arrival of Captain Cook. It is not surprising, therefore, to see that these two ethnic groups, Caucasians and Hawaiians or Part Hawaiians, display the largest proportion of inconsistent responses.

The impact of the way the item is phrased is magnified even further in the presence of the phenomenon of social desirability. It has been shown that, among some minorities, ethnicity is often denied or repressed (Corenblum and Anis, 1993), and Polish Americans (Malewska-Peyre, 1994), while other minorities are experiencing a surge in ethnic pride. With both categories, social desirability could play a big role in the selection of the ethnic group one identifies with, casting further doubt about the integrity of ethnic taxonomy (Weinstein, Manicas, and Leon, 1990). In Hawaii for example, it has become increasingly socially desirable to be "local." Given the opportunity —in the form of an increase in ethnic categories to choose from— some students in Hawaii will report an ethnic minority, albeit constituting a small part of their blood mix. That is perhaps why previous findings report that consistency in responses is highest when students are asked to respond Yes or No, and decreases with the increase in categories to choose from (Walsh, 1967). However, it is obvious that, as the number of categories decrease, so does the value of the report.

Summary and Conclusion

In an environment of increasing ethnic mix, data based on self-reported ethnicity appear to be marred with inconsistency. In addition to possible carelessness on the part of the students, this study suggests that this inconsistency seems to be aggravated by two factors:

1. The phrasing of the items: the fewer the levels of detail the smaller the discrepancy.
2. The extent of blood mix: the more ethnically mixed the student the larger the discrepancy.

At its February 1999 meeting, the National Post-Secondary Education Council panel on race and ethnicity recommended a large increase in the number of ethnic categories, including several combinations of ethnic categories (Electronic AIR, March 1999). The panel also recommended that a unified way of asking for ethnicity be used, asking the students to select "one or more" of the ethnic categories provided. Although this change is needed in light of the increasing blood mix among peoples, its initial impact will undoubtedly lead to greater discrepancy when compared to previous reporting categories. Also, the proliferation of mixed categories will dilute the usefulness of data resulting from ethnic comparisons. Perhaps the day when analyses by ethnicity will be of no use is not too far away after all.

References

Corenblum, B., and Anis, R. C. (1993). Development of racial identity in minority and majority children: an effect discrepancy model. *Canadian Journal of Behavioral Sciences*, 25 (4), pp. 199-251.

Laing, J., Sawyer, R., and Noble, J. (1987). Accuracy of self-reported activities and accomplishments of college-bound students. ACT Research Report Series 87-6. American College Testing, Iowa.

Malewska-Peyre, H. (1994). Identity crises in migrant's [sic] children and their consequences., *Polish Sociological Review*, 2 (106), pp. 125-132.

Porst, R., and Zeifang, K. (1987). A description of the German General Social Survey test-retest study and a report on the stability of the sociodemographic variables. *Sociology Methods and Research*, 15 (3), pp. 177-218.

Takalkar, P., Waugh, G., and Micceri, T. (1993). A search for TRUTH in student responses to selected survey items. Paper presented at the AIR Forum, Chicago.

Walsh, W. B. (1967). Validity of self-reports. *Journal of Counseling Psychology*, 14, pp. 18-23.

Weinstein, M. G., Manicas, P.T., and Leon, J. J. (1990). The Portuguese and Haoles of Hawaii. *American Sociological Review*, 55 (2), pp. 305-308.